1 Chapter 1: Diagonalization of matrices

1.1 Definitions

Let E be an $n-$dimensional space vector over a field K, where $K = \mathbb{R}$ or $\mathbb{C}$.
$dimE = n$, B a basis of E. Let $f : E \longrightarrow E$ a linear application (endomorphism of E), A the square matrix $(n \times n)$ associated with f: $A = \mathbb{M}_B(f) = (a_{ij})$.

1.1.1 Definition 1. Characteristic Polynomial of a Matrix

If A is an $n \times n$ matrix, the **characteristic polynomial** $P(\lambda)$ of A is defined by:

$$P(\lambda) = det(A - \lambda I_n)$$

1.1.2 Definition 2. Eigenvalues and Eigenvectors

If A is $n \times n$ matrix, a number λ is called an eigenvalue of A if there is $V \in E$ such that:

$$AV = \lambda V$$

In this case, V is called an eigenvector of A corresponding to the eigenvalue λ.

Example. If $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ and $V = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ then $AV = \begin{bmatrix} 20 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 5 \\ 1 \end{bmatrix} = 4V$
So $\lambda = 4$ is an eigenvalue of A with corresponding eigenvector V.

Theorem. Let A be an $n \times n$ matrix.

1. The eigenvalues λ of A are the roots of the characteristic polynomial $P(\lambda)$ of A.
$$P(\lambda) = 0$$
2. The $\lambda-$ eigenvectors X are the nonzero solutions to the homogeneous system
$$(A - \lambda I)X = 0$$

1.1.3 Definition 3.

Let A be $n \times n$ matrix and λ an eigenvalue of the matrix A. The set

$$E(\lambda) = \{V \in E, AV = \lambda V\}$$

is called the **eigenspace** of A associated to the eigenvalue λ in which $E(\lambda)$ is vector sub-space of E. Its dimension $(dimE(\lambda))$ is called the the geometric multiplicity of λ.

1.1.4 Definition 4. Similarity and Diagonalization

If A, B are two $n \times n$ matrices, then they are **similar** if and only if there exists an invertible matrix P such that:

$$A = P^{-1}BP$$

1.1.5 Definition 5. Trace of a matrix

If $A = (a_{ij})$ is an $n \times n$ matrix, then the trace of A is

$$trace(A) = tr(A) = \sum_{i=1}^{n} a_{ij}$$

Lemma. Properties of a trace For $n \times n$ matrices A and B, and any $k \in \mathbb{R}$,

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(kA) = k.tr(A)$
3. $tr(AB) = tr(BA)$

Theorem. Properties of similar matrices If A and B are $n \times n$ matrices and A, B are similar, then

1. $det(A) = det(B)$
2. $rank(A) = rank(B)$
3. $tr(A) = tr(B)$
4. $P_A(\lambda) = P_B(\lambda)$
5. A and B have the same eigenvalues.

Proof. **1.** We have $B = P^{-1}AP$, then $det(B) = det(P^{-1}AP) = det(A)$

4. $P_B(\lambda) = det(B - \lambda I_n) = det(P^{-1}AP - P^{-1}\lambda P) = det[P^{-1}(A - \lambda I_n)P] = det(P^{-1}) \times det(A - \lambda I_n) \times det(P)$

□

1.1.6 Definition 6. Digonalizable

Let A be an $n \times n$ matrix. Then A is said to be **diagonalizable** if there exists an invetible matrix P such that

$$P^{-1}AP = D$$

where D is a diagonal matrix.

Proposition. Let λ_1 and λ_2 be two distinct eigenvalues ($\lambda_1 \neq \lambda_2$) of A, then

$$E(\lambda_1) \cap E(\lambda_2) = \{0\}$$

Proof. If $V \in E(\lambda_1) \cap E(\lambda_2)$, then $AV = \lambda_1 V = \lambda_2 V$ i.e. $(\lambda_1 - \lambda_2)V = 0$. Since $\lambda_1 \neq \lambda_2$, then we have $V = 0$ □

1.1.7 Definition 7. Diagonalization

A square $n \times n$ matrix A is **diagonalizable** if A is similar to a diagonal matrix, i.e.

$$A = PDP^{-1}$$

for a diagonal matrix D and an invertible matrix P.

Proposition. Let A be an $n \times n$ matrix. We suppose that $P(\lambda)$ have k distinct roots $\lambda_1, \lambda_2, ..., \lambda_k$. If $E = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_k)$, then A is diagonalizable.

Proof. For $i = 1, 2, ..., k$, we choose the basis B_i of $E(\lambda_i)$. The basis $B' = \cup_{i=1}^{i=k} B_i$ of E consists of the eigenvectors of A associated with the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$, then the matrix $D = \mathbb{M}_{B'}(f)$ is diagonal. □

Examples Find the characteristic polynomial, eigenvalues and eigenvectors of the matrices:

1. $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$
2. $A = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{bmatrix}$

Solution.

1. $P(\lambda) = (\lambda - 4)(\lambda + 2)$
   $\lambda_1 = -2$ and $\lambda_2 = 4$
   $V_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $V_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

2. $P(\lambda) = -\lambda(\lambda - 1)(\lambda - 2)$
   $\lambda_1 = 0$,$\lambda_2 = 4$ and $\lambda_3 = 2$
   $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $V_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

1.2 Sufficient condition for a matrix to be diagonalizable

Proposition. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. We have $P(\lambda) = (-1^n)(\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$, where $\lambda_1, \lambda_2, ..., \lambda_n$ n distinct eigenvalues of A and $V_1, V_2, ..., V_n$ the n eigenvectors associated with λ_i.
$AV_1 = \lambda_1 V_1$
$AV_2 = \lambda_2 V_2$
.
.
.
$AV_n = \lambda_n V_n$
We can prove that $B' = (V_1, V_2, ..., V_n)$ is a basis of E by induction:
We prove that the set $(V_1, V_2, V_3, ..., V_{k+1})$ is linearly independent of E.

$$\alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_k V_k + \alpha_{k+1} V_{k+1} = 0 \tag{1}$$

We have $A(\alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_k V_k + \alpha_{k+1} V_{k+1}) = 0$, then
$\alpha_1 AV_1 + \alpha_2 AV_2 + ... + \alpha_k AV_k + \alpha_{k+1} AV_{k+1} = 0$

$$\alpha_1 \lambda_1 V_1 + \alpha_2 \lambda_2 V_2 + ... + \alpha_k \lambda_k V_k + \alpha_{k+1} \lambda_{k+1} V_{k+1} \tag{2}$$

From $(2) - \lambda_{k+1}(1)$:
$(\lambda_1 - \lambda_{k+1})\alpha_1 V_1 + (\lambda_2 - \lambda_{k+1})\alpha_2 V_2 + ... + (\lambda_k - \lambda_{k+1})\alpha_k V_k = 0$
Since the set $(V_1, V_2, ..., V_k)$ is linearly independent of E by induction hypothesis, then $(\lambda_1 - \lambda_{k+1})\alpha_1 = (\lambda_2 - \lambda_{k+1})\alpha_2 = ... = (\lambda_k - \lambda_{k+1})\alpha_k = 0$ (because λ_k are distinct).
Therefore $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$
By (1) we have $\alpha_{k+1} V_{k+1} = 0$, then $\alpha_{k+1} = 0$ □

1.3 Necessary and sufficient condition for diagonalizability

Proposition 1. Let A be an $n \times n$ matrix, then

$$dim(E(\lambda_1)) \leq m_1$$

where λ_1 is an eigenvalue of A multiplicity m_1.

Proof. Let $(e_1, e_2, ..., e_r)$ the basis of $E(\lambda_1)$, then we can find the basis $B = (e_1, e_2, ..., e_r, e_{r+1}, ..., e_n)$ of E.
The matrix A is similar of the matrix A' of the form

$$A' = \left(\begin{array}{cccc|c} \lambda_1 & & & & \\ & \lambda_1 & & & \\ & & \ddots & & A_1 \\ & & & \lambda_1 & \\ \hline & & 0 & & A_2 \end{array}\right)$$

$$P(\lambda) = det(A - \lambda I_n) = \left|\begin{array}{cccc|c} \lambda_1 - \lambda & & & & \\ & \lambda_1 - \lambda & & & \\ & & \ddots & & A_1 \\ & & & \lambda_1 - \lambda & \\ \hline & & 0 & & A_2 - \lambda I_{n-r} \end{array}\right|$$

$$= (\lambda_1 - \lambda)^r det(A_2 - \lambda I_{n-r})$$

Then $m \geq r$, where $r = dimE(\lambda_1)$ □

Proposition 2. Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if:

1. $P(\lambda)$ is factored.
2. For each eigenvalue λ_i of A, $dim(E(\lambda_i)$ is equal to the multiplicity of λ_i i.e.

$$dimE(\lambda_i) = m_i, i = 1, ..., k$$

Proof. By induction, the sub-spaces $E(\lambda_i)$, $i = 1, ..., j$, verify

$$E = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_k)$$

for $j = 1, ..., k$
Denote $S_j = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_j)$
It is sufficient to demonstrate that $S_j \cap E(\lambda_{j+1}) = \{0\}$
Let $V \in S_j \cap E(\lambda_{j+1})$, then

$$\begin{cases} V = V_1 + V_2 + ... + V_j \\ \text{and} \\ AV = \lambda_{j+1} V \end{cases} \tag{3}$$

For (3), we have $AV = AV_1 + AV_2 + ... + AV_j$, then

$$\lambda_{j+1} V = \lambda_1 V_1 + \lambda_2 V_2 + ... + \lambda_j V_j \tag{4}$$

For $(4) - \lambda_{j+1}(3)$, we have

$$0 = (\lambda_1 - \lambda_{j+1}) V_1 + (\lambda_2 - \lambda_{j+1}) V_2 + ... + (\lambda_j - \lambda_{j+1}) V_j$$

Using induction hypothesis, we get $V_1 = V_2 = ... = V_j = 0$
Since $\sum_{i=1}^{n} dim E(\lambda_i) = \sum_{i=1}^{n} m_i = n$, we see that $E = \oplus_{i=1}^{k} E(\lambda_i)$. Then A is diagonalizable and we write:

$$D = \begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_k & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_k \end{bmatrix}$$

□

Examples.

1. $A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}$

   $P(\lambda) = -\lambda(\lambda - 1)^2$

   $P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$

   $E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$

   $E(\lambda_2) = E(1) =< V_2, V_3 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$, $V_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $dimE(\lambda_2) = 1 = m_2 = 2$.

   Then the matrix A is diagonalizable.

2. $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{bmatrix}$

   $P(\lambda) = -\lambda(\lambda - 1)^2$

   $P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$

   $E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$

   $E(\lambda_2) = E(1) =< V_2 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $dimE(\lambda_2) = 1 \neq m_2 = 2$

   Then the matrix A isn't diagonalizable.

2 Chapter 2: Triangulability of matrices

Example 1. Consider the matrix $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{bmatrix}$, then

$P(\lambda) = -\lambda(\lambda - 1)^2$

$P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$

$E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$

$E(\lambda_2) = E(1) =< V_2 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $dimE(\lambda_2) = 1 \neq m_2 = 2$

Then the matrix A isn't diagonalizable.

What to do if matrix A is not diagonalizable?
Therefore, we use triangulation:

2.1 Proposition

Let $f : E \to F$ a linear map and A the matrix of f, we suppose the characteristic polynomial $P(\lambda)$ of f (or A) is factored in $K[\lambda]$. Then f (or A) is triangulable.

Proof. By induction over $dimE$: the result is true for the space of dimension 1. Suppose they are true for spaces of dimension $\leq n - 1$ and let E be a space of dimension n.
Let $P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$ in $K[\lambda]$, ($K = \mathbb{R}$ or $\mathbb{C}$).
We suppose that the eigenvalues λ_i are not necessarily distinct. We denote V_1, an eigenvector associated with λ_1 (i.e $f(V_1) = \lambda_1 V_1$).
By the incomplete basis theorem, there exists a basis B' of E where $B' = (V_1, e_2, e_3, ..., e_n)$ then the matrix A' has the form

$$A' = M_{B'}(f) = \begin{bmatrix} \lambda_1 & a_{12} & . & . & . & a_{1n} \\ 0 & a_{22} & & & & . \\ 0 & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & a_{n2} & . & . & . & a_{nn} \end{bmatrix}$$

The family $B_1 = (e_2, ..., e_n)$ is a basis of the subspace $F =< e_2, ..., e_n >$ of E. We denote $g : F \to F$, the linear map such that the associated matrix is

$$A_1 = \begin{bmatrix} a_{12} & . & . & . & a_{1n} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{n2} & . & . & . & a_{nn} \end{bmatrix} = M_{B_1}(g)$$

Then $P(\lambda) = (\lambda_1 - \lambda) \times det(A_1 - \lambda I_{n-1})$
i.e. $P(\lambda)$ is factored and since $dimF = n - 1$, by induction hypothesis, there exists a basis $B_2 = (V_2, ..., V_n)$ of F such that $M_{B_2}(g)$ is upper triangular. We get

$$M_{B'=(V_1,V_2,...,V_n)}(f) = \begin{bmatrix} \boxed{\lambda_1} & a_{12} & . & . & . & a_{1n} \\ & \lambda_2 & . & . & . & . \\ & & & . & . & . \\ & & & & . & . \\ & & & & & \lambda_n \end{bmatrix}$$

□

Remark.
1/ If A is triangulable, the diagonal of the matrix $T = M_{B'}(f)$ are the eigenvalues of A.
2/ All matrix of $A \in M_n(\mathbb{C})$ is triangulable.

Corollary.
$tr(A) = \sum_i \lambda_i$
$det(A) = \prod_i \lambda_i$

Remark.
We can triangulate the matrix A of Example 1.

We consider the basis B' of E where
$$\begin{cases} V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = e_1 + e_2 + e_3 \\ V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = e_1 + 3e_2 + 2e_3 \\ V_3 = e_1 \end{cases}$$

Because $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 2 & 0 \end{vmatrix} = 2 - 3 = -1 \neq 0$

And $\begin{cases} e_1 = V_3 \\ e_2 = -2V_1 + V_2 + V_3 \\ e_3 = 3V_1 - V_2 - 2V_3 \end{cases}$

Then $T = M_{B'}(f) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = P^{-1}AP$

Where $\begin{cases} f(V_1) = \lambda_1 V_1 = 0 \\ f(V_2) = \lambda_2 V_2 = V_2 \\ f(V_3) = f(e_1) = e_1 + 2e_2 + e_3 = -V_1 + V_2 + V_3 \end{cases}$

Finally, $T = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ is the upper triangular matrix,

$P = (V_1V_2V_3) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 2 & 0 \end{bmatrix}$ and $P^{-1} = (e_1e_2e_3) = \begin{bmatrix} 0 & -2 & 3 \\ 0 & 1 & -1 \\ 1 & 1 & -2 \end{bmatrix}$

2.2 Annihilating polynomials

Let E a vector space over K and $R \in K[\lambda]$
$R(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_2\lambda_2 + a_1\lambda^1 + a_0\lambda^0$
If $f \in End_K(E)$, we denote $R(f)$, the linear map of E defined by
$R(f) = a_nf^n + a_{n-1}f^{n-1} + ... + a_2f^2 + a_1f^1 + a_0id$
or $R(A)$ the matrix
$R(A) = a_2A^n + a_{n-1}A^{n-1} + ... + a_2A^2 + a_1A^1 + a_0I_n$
Where $f^k = \underbrace{f \circ f \circ ... \circ f}_{\text{k times}}$

Remark.
We have $P(f) \circ Q(f) = Q(f) \circ P(f)$.

2.2.1 Definition.

Let $f \in End_K(E)$, the polynomial $R \in K[\lambda]$ is called annihilating polynomial of f (or A), if

$$R(f) = 0$$

or

$$R(A) = 0$$

.

2.3 Cayley-Hamilton theorem

Let $f \in End_k(E)$ and $P(\lambda)$ the characteristic polynomial of f (or A).
Then

$$P(f) = 0$$

(or $P(A) = 0$). i.e $P(\lambda)$ annihilates f (or A).

Proof. We suppose $K = \mathbb{C}$, in this case f (or A) is triangulable.
Let $B' = (V_1, V_2, .., V_n)$, a basis of E such that

$$M_{B'}(f) = \begin{pmatrix} \lambda_1 & a_{12} & . & . & . & a_{1n} \\ & \lambda_2 & a_{23} & . & . & a_{2n} \\ & & . & & & . \\ & & & & . & . \\ & & & & & \lambda_n \end{pmatrix} = T \text{ is an upper triangular matrix}$$

We have $f(V_1) = \lambda_1 V_1 \Rightarrow (\lambda_1 id - f)(V_1) = 0$ and
$P(\lambda) = det(T - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda)$
Then $P(f) = (\lambda_1 id - f) \circ ... \circ (\lambda_n id - f)$ and
$P(f)(V_1) = (\lambda_2 id - f) \circ ... \circ (\lambda_n id - f) \circ (\lambda_1 id - f)(V_1) = 0$. Therefore, $P(f)(V_1) = 0$
$P(f)(V_2) = (\lambda_3 id - f) \circ ... \circ (\lambda_n id - f) \circ (\lambda_1 id - f) \circ (\lambda_2 id - f)(V_2) = (\lambda_3 id - f) \circ ... \circ (\lambda_n id - f) 0 (\lambda_1 id - f)(-a_{12} V_1) = 0$. Therefore, $P(f)(V_2) = 0$
We can similarly show that $P(f)(V_3) = 0$
By induction, we find $P(f)(V_i) = 0, \forall i = 1, ..., n$. Finally, $P(f) = 0$.

□

Example.

$$A = \begin{bmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 6 & 1 & 1 \end{bmatrix}$$

$P(\lambda) = det(A - \lambda I_3) = (2 - \lambda)(1 - \lambda)^2 = -\lambda^3 + 4\lambda^2 - 5\lambda + 2$
Since $det(A) = P(0) = 2 \neq 0$, A is invertible.
By the Cayley-Hamilton theorem, we have $P(A) = 0$
i.e $-A^3 + 4A^2 - 5A + 2I_3 = 0$. Then $-A^3 + 4A^2 - 5A = -2I_3 \Rightarrow$
$A[-A^2 + 4A - 5I_3] = -2I_3 \Rightarrow A[\frac{1}{2}A^2 - 2A + \frac{5}{2}I_3] = I_3$
Therefore,

$$A^{-1} = \frac{1}{2}A^2 - 2A + \frac{5}{2}I_3$$

2.4 Proposition

Let $S(\lambda)$ a annihilating polynomial of f $[S(f) = 0]$.
All eigenvalue λ_1 of f (of A) is a root of $S(\lambda)[S(\lambda_1) = 0]$.

Proof. If λ_1 is a $V.P, f(V) = \lambda_1 V$
or $S(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0$
$S(f) = a_n f^n + a_{n-1}f^{n-1} + ... + a_1 f + a_0 id = 0$
Therefore $a_n f^n(V) + a_{n-1}f^{n-1}(V) + ... + a_1 \underbrace{f(V)}_{\lambda V} + a_0 id(V) = 0$
$\Rightarrow a_n\lambda^n V + a_{n-1}\lambda^{n-1}V + ... + a_1\lambda V + a_0 V = 0$
$\underbrace{(a_n\lambda_1^n + a_{n-1}\lambda_1^{n-1} + ... + a_1\lambda_1 + a_0)}_{S(\lambda_1)} V = 0.$
Consequently, $[V \neq 0] \Rightarrow S(\lambda_1) = 0$
i.e λ is a root of $S(\lambda)$. □

2.5 Proposition

Let $f \in End(E)$ and $P(\lambda)$ the characteristic polynomial of f i.e

$$P(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_p)^{m_p}$$

If f is diagonalizable, then the polynomial $Q(\lambda) = (\lambda - \lambda_1)...(\lambda - \lambda_p)$ annihilates $f[Q(f) = 0]$.

Proof. If f is diagonalizable, there exists a basis $B' = (V_1, V_2, ..., V_n)$ formed of eigenvectors.
Let $\lambda_1, \lambda_2, ..., \lambda_p$ be the eigenvalues of A. For all $V_i \in B'i = \overline{1, n}$, there exists $\lambda_i \;\; 1 \leq j \leq p$, such that $f(V_1) = \lambda_j V_i$
i.e $(f - \lambda_j id)(V_i = 0)$
$Q(f) = (f - \lambda_1 id) \circ (f - \lambda_2 id) \circ ... \circ (f - \lambda_p id)$
$Q(f)(V_i) = [(f - \lambda_1 id) \circ (f - \lambda_2 id) \circ ... \circ (f - \lambda_p id)](V_i) = (f - \lambda_1 id) \circ ... \circ \underbrace{(f - \lambda_j id)(V_i)}_{0} = 0$ □

2.6 Minimal polynomial

2.6.1 Definition.

We call the **minimal polynomial** of f (or of A) denoted $Q(f)$ (or $Q(A)$), the normalized annihilating polynomial of f (or of A) of the smallest degree.

$$Q(f) = 0 \text{ or } Q(A) = 0$$

Remark. If $S(\lambda)$ is a multiple of $Q(\lambda)$, then

$$S(\lambda) = Q(\lambda) \times T(\lambda)$$

$$S(f) = Q(f) \circ T(f) = 0$$

i.e $S(\lambda)$ is an annihilating polynomial.
Proposition 1.
The annihilating polynomials of f are the polynomials of the type:

$$S(\lambda) = Q(\lambda) \times T(\lambda)$$

Then $S(\lambda) = Q(\lambda) \times T(\lambda) + R(\lambda)$
$S(f) = R(f) = 0 \;\; R(f) = 0$
i.e R is annihilating and since $d^\circ R(\lambda) < d^\circ Q(\lambda)$. This contradicts the hypothesis that $Q(\lambda)$ is a minimal polynomial. Then $R(\lambda) = 0$.
Remark.
$Q(\lambda)/P(\lambda)$ or $P(\lambda) = Q(\lambda) \times T(\lambda])$
Proposition 2.
The roots of $Q(\lambda)$ are exactly the roots of $P(\lambda)$, i.e the eigenvalues but with a different multiplicity
If

$$P(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_p)^{m_p} \qquad \lambda_i \neq \lambda_j$$

Then

$$Q(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2}...(\lambda - \lambda_p)^{\alpha_p}$$

with $1 \leq \alpha_i \leq m_i, \;\; i = 1, ..., p$

Proof. We know that $P(\lambda) = Q(\lambda)T(\lambda)$, then if λ is a root of $Q(\lambda)$, then it is a root of $P(\lambda)$.
Conversely, let λ a root of $P(\lambda)$ i.e λ is an eigenvalue of A, then λ is a root of $Q(\lambda)$ because $Q(\lambda)$ annihilates A. □

2.6.2 Theorem 1.

The minimal polynomial and characteristic polynomial of f (or A) share the same roots, except for multiplicities.

Examples.

- $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$
We have $P(\lambda) = -(\lambda+1)(\lambda+2)(\lambda-3)$, then $Q(\lambda) = (\lambda+1)(\lambda+2)(\lambda-3)$

- $A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$
We have $P(\lambda) = -(\lambda-1)(\lambda+2)^2$, then there exists two possibilities:

$$Q(\lambda) = (\lambda-1)(\lambda+2)$$
$$Q(\lambda) = (\lambda-1)(\lambda+2)^2$$

2.6.3 Theorem 2.

An endomorphism f (or A) is diagonalizable if and only if the minimal polynomial of f (or A) is factored and has all its simple roots.
i.e

$$Q(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)...(\lambda-\lambda_m)$$

Examples.

- $A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$
We saw that $Q(\lambda) = (\lambda+2)(\lambda-1)$, then A is diagonalizable.

- $A = \begin{bmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$
We have $P(\lambda) = -(\lambda-1)^3$, then $Q(\lambda) = \lambda - 1$ or $(\lambda-1)^2$ or $(\lambda-1)^3$
If $Q(\lambda) = \lambda - 1$, $Q(A) = 0$ or $Q(A) = A - I_3 = \begin{bmatrix} 2 & 2 & -2 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \neq 0$
then A is not diagonalizable.

- $A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

  We have $P(\lambda) = -(\lambda-1)(\lambda-2)^2$, then $Q(\lambda) = (\lambda-1)(\lambda-2)$ or $Q(\lambda) = (\lambda-1)(\lambda-2)^2$

  If $Q(\lambda) = (\lambda-1)(\lambda-2)$, then

  $Q(A) = (A - I_3)(A - 2I_3) = \begin{bmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & . & . \\ . & . & . \\ . & . & . \end{bmatrix} \neq 0$

  Then A is not diagonalizable.

2.7 Kernel decomposition theorem

1. We suppose there exists $P \in K[\lambda]$ of the form $P = S \times T$ with $S, T \in K[\lambda]$ relatively prime, such that $P(f) = 0$. Then $E = kerS(f) \oplus kerT(f)$.

2. We suppose there exists $P \in K[\lambda]$ of the form $P = P_1 \times P_2 \times ... \times P_k$ with $P_1, P_2, ..., P_k \in K[\lambda]$ relatively prime pairwise, such that $P(f) = 0$. Then, $E = kerP_1(f) \oplus ... \oplus kerP_k(f)$.

Proof.

1. We prove that $kerS(f) \cap kerT(f) = \{0\}$

   Let $v \in kerS(f) \cap kerT(f)$

   $S(f)(v) = 0$ and $T(f)(v) = 0$

   Or $P(\lambda) = S(\lambda) \times T(\lambda) \Rightarrow P(f) = S(f) \circ T(f)$, since $S(\lambda) \wedge T(\lambda) = 1$.

   Using Besout theorem, $\exists S_1(\lambda), T_1(\lambda)$ such that $S_1(\lambda) \times S(\lambda) + T_1(\lambda) \times T(\lambda) = 1$

   Therefore, $S_1(f) \circ S(f) + T_1(f) \circ T(f) = id$ and

   $v = id(v) = S_1(f)[\underbrace{S(f)(v)}_{0}] + T_1(f)[\underbrace{T(f)(v)}_{0}]$. Then $v = 0$.

   Let $v \in E$

   $v = id(v) = \underbrace{S_1(f) \circ S(f)(v)}_{V_2 \in kerT(f)} + \underbrace{T_1(f) \circ T(f)(v)}_{V_1 \in kerS(f)}$

   $v_1 \in kerS(f)$

   i.e $S(f)(v_1) = S(f)[T_1(f) \circ T(f)(v)] = T_1(f) \circ \underbrace{S(f) \circ T(f)}_{P(f)=0}(v) = 0$.

   Similarly for v_2, we obtain $v_2 \in kerT(f)$

   i.e $v = v_1 + v_2$ □

2.7.1 Proposition

An endomorphism f (or A) is diagonalizable if and only if the minimal polynomial of f (or A) is factored and has all its simple roots.

Proof. If f is diagonalizable $\Rightarrow Q(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_p)$
If $Q(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_p)$
$Q(\lambda) = P_1 \times P_2 \times ...P_p$ with $P_i = \lambda - \lambda_i \in K[\lambda], i = 1, ..., p$ relatively prime pairwise, such that $Q(f) = 0$
Then $E = kerP_1(f) \oplus ... \oplus kerP_p(f) = ker(f - \lambda_1 id) \oplus ... \oplus ker(f - \lambda_p id) = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E_{\lambda_p}$.
i.e E is the direct sum of the eigenspace $E(\lambda_i), i = 1, ..., p$. Then f (or A) is diagonalizable. □

2.8 Applications

- **Compute the power of the matrix**
  Let A be an $n \times n$ matrix.
  Method 1. Using the formula $A = PDP^{-1}$
  We suppose A is diagonalizable, then $D = P^{-1}AP$, i.e $A = PDP^{-1}$, then

$$A^k = (PDP^{-1})(PDP^{-1})...(PDP^{-1}) = PD^kP^{-1}$$

Or $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, $D^k = \begin{bmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix}$

and it's easy to compute A^k using the following formula $A^k = P \begin{bmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix} P^{-1}$

$$A^n = PD^nP^{-1}$$

Method 2. Using the minimal polynomial $Q(\lambda)$
$\lambda^n = Q(\lambda) \times S(\lambda) + R(\lambda)$. Then $A^n = R(A)$

Example.
Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$
We have $P(\lambda) = (\lambda - 2)(\lambda - 3)$

$E(\lambda_1) = E(2) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$E(\lambda_2) = E(3) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

Therefore, $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} -2 & 1 \\ -1 & -1 \end{bmatrix}$

We obtain $A^k = \begin{bmatrix} 2^{k+1} - 3^k & 2^{k+1} - 2.3^k \\ -2^k + 3^k & -2^k + 2.3^k \end{bmatrix}$

- **Solving a system of recurrence relations**

  Let's illustrate this with an example. This involves determining two sequences (u_n), v_n such that:

$$(1) \begin{cases} u_{n+1} = u_n - v_n \\ v_{n+1} = 2u_n + 4v_n \end{cases} \quad \text{and such that} \begin{cases} u_0 = 2 \\ v_0 = 1 \end{cases}$$

We put $X_n = \begin{bmatrix} u_n \\ v_n \end{bmatrix}$. We can write the system (1):

$$X_{n+1} = AX_n \text{ with } A = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

Hence, by induction

$$X_n = A^n X_0 \text{ with } X_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

We have $\begin{bmatrix} u_n \\ v_n \end{bmatrix} = A^k = \begin{bmatrix} 2^{k+1} - 3^k & 2^{k+1} - 2.3^k \\ -2^k + 3^k & -2^k + 2.3^k \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Finally,

$$\begin{cases} u_n = 3.2^{n+1} - 4.3^n \\ v_n = -3.2^n + 4.3^n \end{cases}$$

- **Solving a first-order linear differential system**

  Let the system $X' = AX$, where $X = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$, $X' = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}$

Example.

$$(\text{I}) \begin{cases} x_1' = x_1 + 2x_2 + -3x_3 \\ x_2' = x_1 + 4x_2 - 5x_3 \\ x_3' = 2x_2 - 2x_3 \end{cases}$$

3 Chapter 3: Nilpotent and exponential matrix

3.1 Nilpotent Matrix

3.1.1 Definition

A nilpotent matrix is a square matrix, there exists an integer m such that

$$N^m = 0$$

.
The integer m is called the nilpotency index. It is the smallest integer such that $N^m = 0$.

Examples.

(a) $A = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix}$
The matrix is nilpotent because by squaring matrix A we get the zero matrix as a result:

$$A^2 = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

(b) $B = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 0 & 3 \\ -1 & 2 & -1 \end{bmatrix}$
Although when raising the matrix to 2 we do not obtain the null matrix:

$$B^2 = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 0 & 3 \\ -1 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 & 1 \\ 3 & 0 & 3 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -6 & 0 & -6 \\ 0 & 0 & 0 \\ 6 & 0 & 6 \end{bmatrix}$$

When calculating the cube of the matrix we do not get the matrix with all the elements equal to zero:

$$\begin{bmatrix} -6 & 0 & -6 \\ 0 & 0 & 0 \\ 6 & 0 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 & 1 \\ 3 & 0 & 3 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

So matrix B is a nilpotent matrix, and since the null matrix is obtained with the third power, its nilpotency index is 3.

3.2 Exponential of a matrix

3.2.1 Definition

If A is a constant $n \times n$ matrix, the matrix exponential e^{At} is given by:

$$e^{At} = I + At + A^2\frac{t^2}{2!} + ... + A^n\frac{t^n}{n!} + ...,$$

where the right-hand side indicates the $n \times n$ matrix whose elements are power series with coefficients given by the entries in the matrices.

Example. The exponential is easiest to compute when A is diagonal. For the matrix $A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$, we calculate

$$A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, A^3 = \begin{bmatrix} -1 & 0 \\ 0 & 8 \end{bmatrix}, ..., A^n = \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}$$

Then we get

$$e^{At} = \sum_{n=0}^{\infty} A^n\frac{t^n}{n!} = \begin{bmatrix} \sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!} & 0 \\ 0 & \sum_{n=0}^{\infty} 2^n\frac{t^n}{n!} \end{bmatrix} = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{bmatrix}$$

Remark. In general, if A is an $n \times n$ matrix with entries $\lambda_1, \lambda_2, ...\lambda_n$, then e^{At} is the diagonal matrix with entries $e^{\lambda_1 t}, e^{\lambda_2 t}, ..., e^{\lambda_n t}$ on the main diagonal.

3.2.2 Theorem 1.

Let A and B be $n \times n$ constant matrices, and $r, s, t \in \mathbb{R}$. Then

(a) $e^{A0} = e^0 = I$

(b) $e^{A(t+s)} = e^{At}e^{As}$

(c) $(e^{At})^{-1} = e^{-At}$

(d) $e^{(A+B)^t} = e^{At}e^{Bt}$ if $AB = BA$

(e) $e^{rIt} = e^{rt}I$

3.2.3 Theorem 2.

If A is an $n \times n$ constant matrix, then the columns of the matrix exponential e^{At} form of a fundamental solution set for the system $x'(t) = Ax(t)$. Therefore, e^{At} is a fundamental matrix for the system, and a general solution is $x(t) = ce^{At}$.

3.3 Exponential of a nilpotent matrix

If A is nilpotent of index m, i.e $A^m = 0$, then

$$e^{At} = I + At + ... + A^{m-1}\frac{t^{k-1}}{(k-1)!}$$

Example. Find the fundamental matrix e^{At} for the system $x' = Ax$, where

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ -2 & -2 & -1 \end{bmatrix}$$

Solution. We find the polynomial of A

$$p(r) = |A - rI| = \begin{vmatrix} 2-r & 1 & 1 \\ 1 & 2-r & 1 \\ -2 & -2 & -1-r \end{vmatrix} = -(r-1)^3$$

Therefore, $r = 1$ is the only eigenvalue of A, so $(A - I)^3 = 0$ and

$$e^{At} = e^t e^{(A-I)t} e^t \{I + (A-I)t + (A-I)^2\frac{t^2}{2}\}.....(1)$$

We calculate

$$A - I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & -2 \end{bmatrix} \text{ and } (A - I)^2 = 0$$

Substitution into (1) gives us

$$e^{At} = e^t I + te^t (A - I) = \begin{bmatrix} e^t + te^t & te^t & te^t \\ te^t & e^t + te^t & te^t \\ -2te^t & -2te^t & e^t - 2te^t \end{bmatrix}$$